

When numbers become words

Citation for published version (APA):

Wilby, K. J. (2019). When numbers become words: Assessors' processing of performance data within OSCEs. [Doctoral Thesis, Maastricht University]. <https://doi.org/10.26481/dis.20190702kw>

Document status and date:

Published: 01/01/2019

DOI:

[10.26481/dis.20190702kw](https://doi.org/10.26481/dis.20190702kw)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Valorisation

Valorisation

1. (Relevance) What is the social relevance of your research results (i.e. in addition to the scientific relevance)?

The results from this research can have potential social advantages for health professional training schools and the development of performance-based assessments.

As described in Chapter 2, we found that culture (as defined by Cultural Dimensions Theory) may influence how assessors interpret and judge communication behaviours. We found that assessors' cultural dimensions of masculinity-femininity and power-distance helped to explain differences in scoring, especially for borderline performers. This is an important finding in a world that is rapidly changing with migration and globalization. Although we are yet to fully understand the impact of culture on assessment, programs must be aware of potential cultural bias and work to ensure assessment practices are fair and provide equal opportunity for all students to be successful.

Our findings from Chapters 3 and 5 support the use of narrative assessment comments in Objective Structured Clinical Examinations (OSCEs) and this may improve defensibility of assessment decisions. In Chapter 3, we found that assessors write narrative comments that can distinguish between good and poor performers and focus on core constructs that are deemed to be fundamental to the patient care process (confidence, adaptability, patient safety, and professionalism). In Chapter 5, we found that scoring based on these comments was just as reliable as scores obtained during the interaction. Having written comments that are known to be reliable and discriminatory of student performance may supplement assessment decisions and provide a better understanding of what actually occurred during the OSCE. This 'record' of performance may therefore aid defensibility of OSCE outcomes, next to providing meaningful feedback to trainees.

Findings from the studies included in this thesis may have advantages with respect to assessor training. Findings from Chapters 2 and 4 showed that assessors' inherent characteristics (cultural dimensions, perspectives) influence how they approach and interpret performance data. The findings from our chapters support a more targeted approach to training by making assessors aware of their own 'approach' to assessment and have the capacity to self-reflect on their characteristics that may be influencing how they interpret and judge performance. If proven to be successful, shifting training accordingly may result in achieving better training outcomes.

2. (Target groups) To whom, in addition to the academic community, are your research results of interest and why?

Results of this thesis may be of interest to many groups and organizations. One major target would be teachers or educators within academic program committees that design performance assessments, including OSCEs. These committees may be interested in how obtaining narrative assessment data from OSCEs may support or enhance assessment approaches in which numeric scores are used to capture and make decisions about student performance. Additionally, our results from Chapters 2 and 4 add to previous research findings by showing that assessor characteristics are important for the interpretation and judgment of student performance in the standardized OSCE format. These results are of interest for faculty developers who must ensure that these issues are paid attention to in assessor training, as well as educators who are responsible for design of assessment systems, as our findings emphasize the need to include multiple assessors and multiple assessor perspectives in our decision making processes.

Secondly, students themselves may be a target group for the results of this research. Participant students were very interested in the narrative comments throughout the research and requested to obtain their own copies in some circumstances. Anecdotally, students seemed to think that comments would help them better understand assessors' scoring decisions and provide individualized performance information that would help to rationalize their grades. Although the perspectives of students were not sought from research within this thesis, it should be a major priority for future study.

Finally, organizations that implement OSCEs for professional licensure or renewal (e.g. Pharmacy Examining Board of Canada, Ontario College of Pharmacists, Pharmaceutical Society of New Zealand) may be interested in the results of this thesis. The findings relating to assessor characteristics, as well as the implications for assessor training and assessment instruments, may be advantageous for these organizations to consider when designing more robust assessment procedures.

3. (Activities/Products) Into which concrete products, services, processes, activities or commercial activities will your results be translated and shaped?

The findings from this thesis can inform the development of new assessor training programs for performance-based assessments of communication skills, including OSCEs. As discussed above, our findings support a targeted approach to assessor training that supports promotion of self-awareness and reflection by assessors, creating awareness of their inherent characteristics and beliefs that may influence the assessment process. Findings from Chapters 2 and 4, for example, suggest assessors could be trained to recognize and acknowledge their own cultural norms or assessment perspectives when observing and interpreting performance. If narratives are to be incorporated into OSCE assessment, as supported by our findings from Chapters 3 and 5, assessors will also need to be trained to write meaningful data within their narrative descriptions of performance. Our results are therefore very relevant to individuals or groups who administer OSCEs and recruit/train assessors.

Findings from this research are already shaping new OSCE assessment methods in different international contexts. The studies took place at Qatar University, where a final cumulative OSCE is maintained as an exit-from-degree requirement for students to complete prior to graduation. The results from Chapters 2 and 3 informed the development of a new communication assessment process, including the collection of narrative data. Having left Qatar and assuming a new post at the University of Otago in New Zealand, I am now implementing similar assessment approaches for our formative and summative OSCEs in my capacity as the Chair of the BPharm Undergraduate Curriculum Committee. These processes include modifications to our communication assessment instruments (as described in Chapter 3), as well as collection of narrative data. Dissemination and future testing of these new processes may result in uptake in other local and international settings.

Three studies from this thesis (Chapters 2, 3, and 5) are available as published manuscripts or available early online for purchase or academic use and have already received citations and over 120 reads on researchgate.net. Chapter 4 is currently under editorial review. The research has also gained interest at international conferences (the Netherlands, New Zealand, Qatar, Canada, Scotland, United Arab Emirates, and Finland), where I have spoken about our findings. . More speaking engagements are being scheduled in other countries in 2019, including Italy. These opportunities will help to disseminate and promote our findings to target groups and audiences.

4. (Innovation) To what degree can your results be called innovative in respect to the existing range of products, services, processes, activities and commercial activities?

To our knowledge, our results are the first to investigate the quality of narrative comments (Chapters 3 and 5) from OSCEs. OSCEs are traditionally scored based on checklists, rating scales, or rubrics and so our results provide an alternative approach to assessing student performance given that narrative assessment comments are reliable and can discriminate between students.

Another area where our results show innovation is the association between assessors' cultural dimensions and scoring of communication skills. To date, research has focused on the influence of culture on communication but yet to show any association with assessment of communication. Our results fill this gap by providing some evidence that assessors' culture may be a factor to consider when designing assessments and forming assessor groups or pairs. However, more research is required to better understand the relationship between culture and assessment and how cultural norms and preferences may influence performance judgements and decisions.

5. (Schedule & Implementation) How will this/these plan(s) for valorization be shaped? What is the schedule, are there risks involved, what market opportunities are there and what are the costs involved?

The valorization of this research can be divided into different streams. First, dissemination of the findings has already begun. To date, three of the chapters are publications available for purchase/downloading online and a fourth is under editorial review. I have also supported these publications with published commentary in journals and scholarly blogs to enhance dissemination of our findings. The work has also been presented at numerous local and international conferences, as described above. Links and key findings have also been shared on social media. The thesis will also be printed as a book and will be publicly available in 2019.

The second area of valorization is the practical uptake of our findings, including the use of narrative comments within OSCEs. As stated above, this is now occurring in two contexts (Qatar and New Zealand). As we gain more experience with using these comments for assessment purposes, we plan to further disseminate our work to the international community, which may increase uptake in other settings. A strong presence at international conferences and on social media platforms over the next 1-2 years should help to expedite this process and facilitate interest and uptake of our results.

Other areas of impact, such as the redesign of assessor training programs, may take longer (within 5 years) to produce. Before new training methods can be implemented, research and testing is required to ensure that training is effective and ultimately enhances the assessment process.